Discussion of "Despiking Acoustic Doppler Velocimeter Data" [1]
by Derek G. Goring and Vladimir I. Nikora

Tony L. Wahl[2]

The writers are commended for their novel approach to the difficult problem of removing spikes from velocity time series recorded by acoustic Doppler velocimeters (ADV's). Their paper has prompted some further investigations into modifications that may make the method even more practicable and robust.

To briefly review, spikes in ADV time series can be caused by many factors, including high turbulence intensities, aerated flows that have undesirable acoustic properties, and phase difference ambiguities that occur when velocities exceed the upper limits of ADV probe velocity ranges. Although spikes can be reduced or eliminated in many cases by adjustment of probe operational parameters, there are some situations in which spikes cannot be entirely avoided. A method for spike detection such as that proposed by the writers is thus necessary.

The phase-space thresholding method set forth by the writers is based on the observation that good ADV data are tightly clumped within an ellipsoid in phase space [three-dimensional (3D) plots of velocity, u, and approximations of the first and second derivatives, Δu and $\Delta^2 u$, respectively]. Spikes are those points plotting outside of elliptical projections of the ellipsoid onto the three principal phase-space planes, u-Δu, Δu-$\Delta^2 u$, and u-$\Delta^2 u$. The first two ellipses are symmetric about both axes, while the third is rotated due to correlation between u and $\Delta^2 u$; positive velocities are associated with negative second derivatives and vice-versa as the velocity tends to move back toward its average value when an extreme is reached. The writers size the ellipsoid and the projected ellipses using the standard deviations of the u, Δu, and $\Delta^2 u$ time series and a Universal Threshold parameter related to the length of the time series. The Universal Threshold requires no tuning, a fundamental advantage of the algorithm. The method is iterative, with spikes identified and replaced, and the algorithm repeated one or more times. Iteration is necessary because the standard deviations, and thus the ellipse boundaries, are sensitive to the spikes and change as spikes are identified and replaced. The writers point out that spike replacement is an arbitrary procedure with many strategies available. Spike replacement is necessary to use the iterative phase-space thresholding method, since derivative and second derivative estimates are needed in subsequent iterations.

Identifying Points Outside of the Ellipsoid

The use of two-dimensional (2D) ellipse projections requires that three comparisons be made to determine if a data point is a spike, one in each principal plane. For each comparison the distance of the data point from the center of the ellipse must be determined and compared to the corresponding distance to a point on the ellipse located along the ray connecting the center of the ellipse and the data point. Spikes are ultimately identified if they are located outside of a

---

[2] Hydraulic Engineer, U. S. Dept. of the Interior, Bureau of Reclamation, Water Resources Research Laboratory, Mail Code D-8560, P.O. Box 25007, Denver, CO 80225-0007. E-mail: twahl@do.usbr.gov
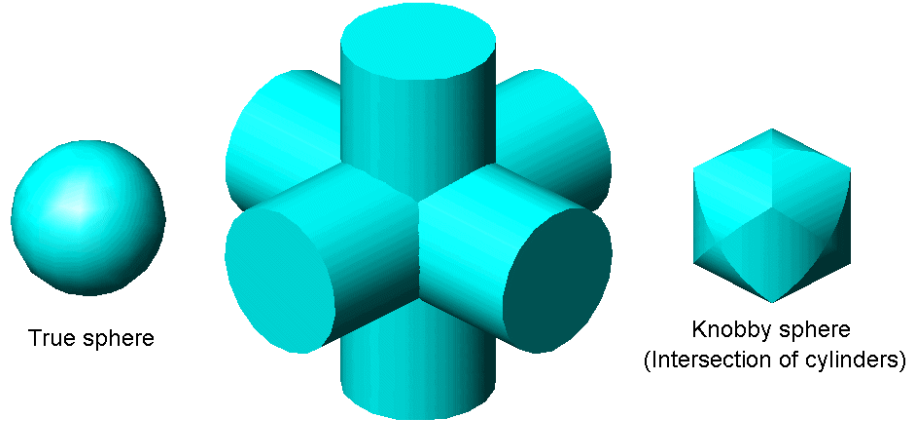
Figure 9. — A true sphere and the object produced by the intersection of orthogonal circular cylinders.

complex three-dimensional shape resembling an ellipsoid, but having knobby projections in each of the eight phase-space quadrants. This object is shown in fig. 9 for the simplified case of three intersecting circular cylinders (i.e., standard deviations of u, Δu, and $\Delta^2 u$ being identical), which produce a knobby sphere. The knobs allow some points that lie outside of the true ellipsoid to be retained as good data.

A single-comparison test using the true ellipsoid shape can identify spikes if the ellipsoid is defined in spherical coordinates, ρ, ϕ, θ (fig. 10)

$$\frac{1}{\rho^2} = \frac{(\sin\phi\cos\theta\cos\alpha + \cos\phi\sin\alpha)^2}{a^2} + \frac{(\sin\phi\cos\theta\sin\alpha - \cos\phi\cos\alpha)^2}{b^2} + \frac{(\sin\phi\sin\theta)^2}{c^2} \qquad (1)$$

where a and b are the major and minor axes of the ellipse projected onto the u-$\Delta^2 u$ plane, as given by Eqs. (10) and (11), and c is the major axis in the Δu-$\Delta^2 u$ projection, identified by the writers as $\lambda_U \sigma_{\Delta u}$. Here, the symbol α indicates the angle of rotation of the ellipsoid in the u-$\Delta^2 u$ plane, given by the writers' Eq. (9). (The writers used the symbol θ for this parameter, which is being used here in connection with the spherical coordinate system). The ellipsoid in Eq. (1) is centered at the origin, requiring one to remove the mean of the velocity time series before testing for spikes.

The test to identify points outside of the ellipsoid is performed in a manner similar to that described above. The position of the u, Δu, $\Delta^2 u$ data point is expressed in spherical coordinates as a distance ρ and the angles ϕ and θ. The distance ρ_e to a point on the surface of the ellipsoid at the same ϕ and θ coordinates is determined from Eq. (1) and ρ and ρ_e are compared to determine if the point is inside or outside the ellipsoid.

Choice of Spike-Detection Threshold

Although applied in 3D phase-space, the method described by the writers is similar to one-dimensional methods for excluding outliers from experimental data. Points lying outside of a

threshold are marked as outliers and removed, or optionally replaced. The threshold used by the writers is $\sigma\sqrt{2\ln(N)}$, in which σ is the standard deviation and N is the number of points. The easily calculated multiplier $\lambda_U = \sqrt{2\ln(N)}$ is the Universal Threshold, described as the theoretical expected maximum value of N independent samples drawn from a standard normal distribution. Thresholds are computed for the three phase-space axes, u, Δu, and $\Delta^2 u$, and these values are used for the ellipse-based tests described earlier.

The writers noted that the Universal Threshold is about 10% larger than expected for $N>1000$. If the threshold were truly the expected maximum value of N independent samples, then we should expect that the performance of the exclusion method would be independent of the value of N. To test this, the expected number of observations rejected can be determined by computing the threshold, $\sqrt{2\ln(N)}$,

Figure 10. — Spherical coordinates and phase space coordinates.

and determining the corresponding rejection probability from a two-tailed standard-normal probability table. The expected number of rejections varies slightly with N, as shown in fig. 11, starting at about 0.5 and decreasing as N increases.

Another common method for excluding outliers from experimental data uses Chauvenet's criterion (Coleman and Steele 1999). For a set of N data, a rejection probability is computed, $p=1/(2N)$, and the Z-statistic for this probability is determined from a two-tailed normal probability table. The Z-statistic for each observation is then computed, $Z=(x-\mu)/\sigma$. Those observations outside of the $\pm Z$ range are excluded. The expected number of observations excluded is always $N/(2N)$, or 0.5, assuming the data are drawn from a standard-normal distribution. This approach has the property we expected for the Universal Threshold; the performance is independent of the size of the sample, and we always stand a coin-flip chance of excluding one observation from a true normally distributed data set. Chauvenet's criterion should be applied only once to a given data set.

For an automated despiking algorithm, an easily calculated threshold is desirable. Chauvenet's criterion can be determined without using lookup tables, since standard-normal probability values can be computed using the error
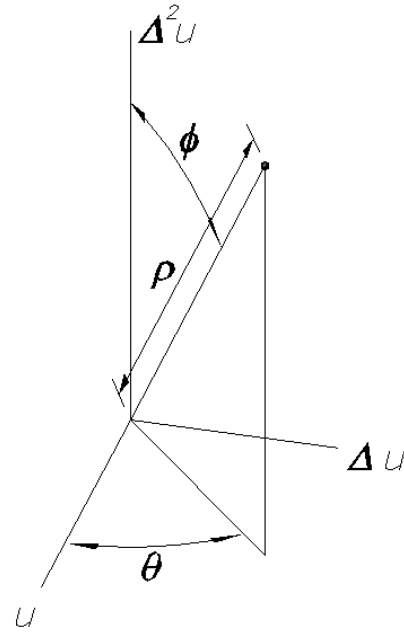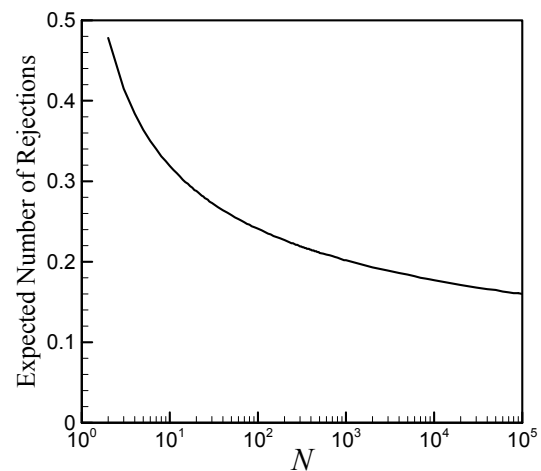

Figure 11. — Expected number of rejections from a normally distributed data set of N elements, using the Universal Threshold.

function, erf(x), determined by series approximation (Kreyszig 1988)

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}}\left[x - \frac{x^3}{1!3} + \frac{x^5}{2!5} - \frac{x^7}{3!7} + \left((-1)^j \frac{x^{2j+1}}{j!(2j+1)}\right)\Bigg|_{j=4}^{j=\infty} \cdots \right] \tag{2}$$

Eq. (2) converges quickly for small values of x, but may require j=30 to 40 terms for x>5. The rejection probability for the two-tailed case for a standardized observation, $Z=(x\text{-}\mu)/\sigma$ is $p = 1 - \text{erf}(Z/\sqrt{2})$ and the Z-statistic for a given rejection probability is $Z = \sqrt{2}\ \text{erfinv}\ (1-p)$. The inverse of the error function, erfinv(1-p), can be determined by Newton's method making use of subroutines that return the error function and its derivative.

Robustness of the Despiking Threshold

Regardless of the threshold chosen, a method that is insensitive to the outliers themselves is desirable. The despiking algorithm presented by the writers uses the mean and standard deviation, the classic estimators for location (general position of the data) and scale (spread of the data), respectively. Rousseeuw (1998) shows that both are very sensitive to outliers, having a breakdown point of 1/N. The breakdown point is the smallest fraction of the values that must be replaced with outliers to carry the estimator over all bounds. One may choose which observations to replace, as well as the magnitude of the outliers, in the most unfavorable way.

When outlier exclusion is based on the mean and standard deviation, a single outlier of extraordinary magnitude can corrupt both parameters and significantly affect performance. The location estimator moves toward the outlier and away from the good data, and the scale estimator explodes, protecting the outliers from rejection. If outlier-exclusion is performed iteratively, the result will improve as outliers are excluded, assuming that the initial iterations exclude some outliers and are not so seriously affected that they exclude non-outlying data.

Outlier-exclusion can be improved by using robust location and scale estimators. A robust estimator of location is the sample median. The median has the largest possible breakdown point, 50%, meaning that one must replace at least half the values to ensure that the estimator will exceed the range of the original data. An extremely robust estimator of scale is based upon the median of the absolute deviations (MAD) from the sample median,

$$S = 1.483\left(\text{median}_{j=1,\ldots,n}\left|x_j - \text{median}_{i=1,\ldots,n}(x_i)\right|\right) = 1.483(\text{MAD}) \tag{3}$$

in which 1.483 is a factor that makes the estimator analogous to the standard deviation, the usual scale parameter of a normal distribution. This estimator also has a breakdown point of 50%.

An outlier-exclusion algorithm described by Rousseeuw (1998) computes the standardized observations, $Z_i = (x_i - T)/S$. The median is used for the location estimator, T, and the scale estimator, S, is based upon the MAD as shown in Eq. (3). The absolute values of the standardized observations are compared to a cutoff value, somewhat arbitrarily chosen as 2.5, and observations above the cutoff are outliers. The cutoff could also be established using the Universal Threshold or Chauvenet's criterion. Regardless, the algorithm need only be applied

once, as the estimators of location and scale will change only slightly after removal of the outliers.

Application to Multidimensional Velocity Time Series

The phase-space threshold despiking method is described by the writers for application to the time series of a single component of velocity. Most ADV probes report three-dimensional velocities (or 2D in some cases). The reported orthogonal velocity components are interdependent, since the instrument actually measures beam velocities along the bi-static axes of the sending and receiving acoustic elements and converts these to orthogonal velocity components via multiplication by a transformation matrix (Kraus et al. 1994). Factors that produce spikes often affect only one of the individual beam velocities (e.g., overranging caused by a large turbulent velocity burst aligned with one of the bi-static axes, or a single air bubble corrupting one acoustic beam), but after multiplication by the transformation matrix, the single affected beam velocity taints all three of the orthogonal velocity components. As a result, a conservative filtering algorithm should eliminate and/or replace associated data points in all three time series when a spike is identified in any one time series, even if the data in the other time series do not *appear* to be spikes.

Robust Phase-Space Despiking Algorithm

A despiking algorithm based on the writers' phase-space thresholding method and incorporating the concepts discussed above was developed and added to the WinADV computer program, a freely distributed software packaged used for post-processing and analysis of ADV data (Wahl 2000). WinADV was developed primarily for use by the Bureau of Reclamation's Water Resources Research Laboratory, where ADV's are used to collect laboratory and field data around model and prototype hydraulic structures and in reservoirs and natural channels. The primary objective of these measurements is usually to define average velocities and turbulence parameters in two or three dimensions.

The modified algorithm incorporates the following features to create a robust, conservative despiking method whose performance is independent of the length of the ADV time series.

- The median and MAD are used as location and scale estimators,

- Chauvenet's criterion is used to define the rejection probability and exclusion thresholds (principal axes of the ellipsoid in phase space),

- The exclusion test uses the true ellipsoid shaped defined in spherical coordinates, and

- Despiking is performed on each of the available velocity components (usually three), and all associated data are removed when a spike is detected in any one of the time series.

Before despiking, the median of each velocity time series is removed. Alternatively, one could high-pass filter the data as described by the writers. Spike replacement is not used, since the algorithm is no longer iterative and does not require replacement of spikes. For most applications of ADV's at the Bureau of Reclamation, a continuous time series is not needed, and a data set that does not contain synthesized or interpolated data is desirable.

Example Application

The method described above was applied to the two data sets examined by the writers. The results are compared in Table 1 to the writers' method. The modified method identifies more spikes than the writers' method in three different ways:

- Excluding data within the knobs of the "knobby" ellipsoid,

- Reducing the size of the ellipsoid by using the MAD as the scale estimator and Chauvenet's criterion to define the rejection probability and exclusion thresholds, and

- Examining all velocity components and rejecting associated data in all time series when a spike is detected in any one component.

Analysis

The modifications to the writers' method each identify more spikes, with the greatest effects in these examples being caused by the reduction of the exclusion thresholds and the examination of all three associated time series. The number of additional spikes located in the knobs of the knobby ellipsoid was relatively small.

In the contaminated data set, examination of the v and w velocity components added about 25% more spikes than were detected in the u component, and of these, about two-thirds came from the

Table 1. — Comparison of despiking methods.

	Contaminated record	Clean record
Length of time series	30,000 samples	30,000 samples
Goring-Nikora method (Universal Threshold)		
Standard deviation of u before despiking	23.52 cm/s	8.058 cm/s
Spikes identified in u in first iteration	858	14
Standard deviation of u after one iteration (spikes removed, but not replaced)	14.68 cm/s	8.050 cm/s
Standard deviation of u after four iterations (spikes replaced using cubic fit)	13.78 cm/s	. . .
Effect of 3-D ellipsoid test (Eq. 12) and identifying spikes in all velocity components		
Spikes in u in first iteration	880	20
Additional spikes in v and w time series	229	160
Robust despiking algorithm using MAD, Chauvenet's criterion, 3-D ellipsoid test, and all velocity components		
1.483(MAD) of u before despiking	13.55 cm/s	7.904 cm/s
Spikes identified in u time series	2149	49
Additional spikes in v and w time series	474	415
1.483(MAD) of u after despiking	12.89 cm/s	7.845 cm/s
Standard deviation of u after despiking	13.19 cm/s	7.994 cm/s

w component. In the clean data set, the *v* and *w* components added about 8 times more spikes than were found in the *u* component alone, and about 85% were due to the *w* component. These results were similar, regardless of which threshold was used (the Universal Threshold or the MAD and Chauvenet's criterion). The relatively large number of additional spikes found in the *w* component may be an anomaly of these two data sets; the writers reported that in their experience with down-looking ADV probes the vertical component (*w* in these cases) usually contains far fewer spikes than the horizontal components.

Using the MAD and Chauvenet's criterion to define the rejection thresholds significantly increases the number of spikes in comparison to the first iteration of the writers' method, although subsequent iterations would identify additional spikes. The advantages of the approach utilizing the MAD are the robustness and insensitivity to the spikes themselves, and the fact that despiking need not be done iteratively.

The methods can also be evaluated by comparing the standard deviations and median absolute deviations before and after despiking, shown in Table 1. In the clean record, the standard deviations and the values of 1.483(MAD) are similar before and after despiking. This is expected, since the outliers in this data set are few in number and not extreme. In the contaminated record (which the writers aptly describe as pathological), the standard deviation of the *u* time series is reduced dramatically from 23.52 cm/s to 14.68 cm/s by the first iteration of the writers' method. Four iterations of despiking and spike replacement reduce the standard deviation to a final value of 13.78 cm/s. In comparison, the value of 1.483(MAD) is 13.55 cm/s before despiking, comparable to the standard deviation *after* despiking. The robust despiking method reduces the value modestly to 1.483(MAD)=12.89 cm/s, demonstrating that this scale estimator is relatively insensitive to spikes and is a reasonable approximation to the standard deviation. The final standard deviation of *u* after applying the robust despiking method is 13.19 cm/s.

References

Coleman, Hugh W., and Steele, W. Glenn (1999). *Experimentation and Uncertainty Analysis for Engineers*, 2nd ed., John Wiley & Sons, Inc., New York, NY, pp. 34-37.

Kraus, Nicholas C., Lohrmann, A., and Cabrera, R. (1994) "New Acoustic Meter for Measuring 3D Laboratory Flows," *Journal of Hydraulic Engineering*, Vol. 120, No. 3, pp. 406-412.

Kreyszig, Erwin (1988). *Advanced Engineering Mathematics*, 6th ed., John Wiley & Sons, New York, NY, pg. A78.

Rousseeuw, Peter J. (1998). "Robust Estimation and Identifying Outliers" in *Handbook of Statistical Methods for Engineers and Scientists*, 2nd ed., Harrison M. Wadsworth, Jr., editor. McGraw-Hill, New York, NY, pp. 17.1-17.15.

Wahl, Tony L., 2000, "Analyzing ADV Data Using WinADV". 2000 Joint Conference on Water Resources Engineering and Water Resources Planning & Management, American Society of Civil Engineers, July 30 - August 2, 2000, Minneapolis, Minnesota. www.usbr.gov/wrrl/twahl/winadv